VERIFICATION

I, Mieko Usuyama of MIZOI International Patent Firm, NTA Ofuna bldg. 3F, 17-10, Ofuna 2-chome, Kamakura-shi, Kanagawa-ken, 247-0056 Japan,

hereby verify that I am the translator of the Japanese Patent Application No. HEI 9-354754 filed on December 24, 1997 for claiming priority of PCT/JP98/05513 and verify that the following is a true and correct translation to the best of my knowledge and belief.

Dated April 5, 2002

Mieko Usuyama

English Translation of Hei 9-354754

[Name of the Document] APPLICATION FOR PATENT

[Arrangement Number] BP508565

[Filing Date] December 24, 1997

[To] The Commissioner of the Patent Office

[International Patent Classification] G10L 9/14

[Title of the Invention] A Method for Speech Coding, Method for Speech Decoding and their Apparatuses

[Number of Claims] 12

[Inventor]

[Address or Domicile] c/o Mitsubishi Electric Engineering Co., Ltd.

6-2, Otemachi 2-chome, Chiyoda-ku, Tokyo

[Name] Tadashi Yamaura

[Applicant]

[Identification Number] 000006013

[Name] Mitsubishi Denki Kabushiki Kaisha

[Agent]

[Identification Number] 100102439

[Patent Attorney]

[Name] Kaneo Miyata

[Appointed Agent]

[Identification Number] 100103894

[Patent Attorney]

[Name] Takeshi Ieiri

[Appointed Agent]

[Identification Number]	100091029	
[Patent Attorney]		
[Name]	Mamoru Ueda	

[Appointed Agent]

[Identification Number]	100092462	
[Patent Attorney]		
[Name]	Yahei Takase	

[Indication of Application Fee]

[Register Number]	011394
[Amount of Payment]	21,000

[List of Documents Attached]

[Document Name]	Specification	1
[Document Name]	Drawings	1
[Document Name]	Abstract	1

[General Power of Attorney Number] 9704079

[Name of the Document] Specification

[Title of the Invention] A Method for Speech Coding, Method for Speech Decoding and their Apparatuses

[Claims]

[Claim 1] A speech coding method according to a code-excited linear prediction (CELP) speech coding method, comprising:

evaluating a noise level of a speech in a concerning coding period by using a code or coding result of at least one of spectrum information, power information, and pitch information; and

selecting a various excitation codebook based on an evaluation result.

[Claim 2] The speech coding method of claim 1, further comprising:

the plurality of excitation codebooks storing time series vectors with various noise levels; and

switching the plurality of excitation codebooks based on the evaluation result of the noise level of the speech.

[Claim 3] The speech coding method of claim 1, further comprising changing a noise level of time series vectors stored in the excitation codebooks based on the evaluation result of the noise level of the speech.

[Claim 4] The speech coding method of claim 3, further comprising:

an excitation codebook storing noise time series vectors; and

generating a low noise time series vector by sampling signal samples in the time series vectors based on the evaluation result of the noise level of the speech.

[Claim 5] The speech coding method of claim 3, further comprising:

a first excitation codebook storing a noise time series vector and a second excitation codebook storing a non-noise time series vector; and

generating a time series vector by adding the time series vector in the first excitation codebook and the time series vector in the second excitation codebook by weighting based on the evaluation result of the noise level of the speech.

[Claim 6] A speech decoding method according to a code-excited linear prediction (CELP) speech decoding method, comprising:

evaluating a noise level of a speech in a concerning decoding period by using a code or decoding result of at least one of spectrum information, power information, and pitch information; and

selecting one of a various excitation codebook based on an evaluation result.

[Claim 7] The speech decoding method of claim 6, further comprising:

the plurality of excitation codebooks storing time series vectors with various noise levels; and

switching the plurality of excitation codebooks based on the evaluation result of the noise level of the speech.

[Claim 8] The speech decoding method of claim 6, further comprising changing a noise level of time series vectors stored in the excitation codebooks based on the evaluation result of the noise level of the speech.

[Claim 9] The speech decoding method of claim 8, further

comprising:

an excitation codebook storing noise time series vectors; and

generating a low noise time series vector by sampling signal samples in the time series vectors based on the evaluation result of the noise level of the speech.

[Claim 10] The speech decoding method of claim 8, further comprising:

a first excitation codebook storing a noise time series vector and a second excitation codebook storing a non-noise time series vector; and

generating a time series vector by adding the code vector in the first excitation codebook and the time series vector in the second excitation codebook by weighting based on the evaluation result of the noise level of the speech.

[Claim 11] A speech coding apparatus, comprising:

a spectrum information encoding means for coding spectrum information of an input speech, and outputting a coded spectrum information as an element of a coding result;

a noise level evaluating means for evaluating a noise level of a speech in a concerning coding period by using a code or coding result of at least one of spectrum information and power information, obtained from the coded spectrum information provided by the spectrum information encoding means, and outputting an evaluation result;

a first excitation codebook storing a plurality of non-noise time series vectors;

a second excitation codebook storing a plurality of noise time series

vectors;

an excitation codebook switching means for switching the first excitation codebook and the second excitation codebook based on the evaluation result by the noise level evaluator;

a weighting-adding means for weighting the time series vectors from the first excitation codebook and second excitation codebook depending on their gains;

a synthesis filter for producing a coded speech based on an excitation signal, which is a weighted time series vector, and the coded spectrum information from the spectrum information encoding means; and

a distance calculating means for calculating a distance between the coded speech and the input speech, searching an excitation code and gain for minimizing the distance, and outputting a result as an excitation code and a gain code as a coding result.

[Claim 12] A speech decoding apparatus, comprising:

a spectrum information decoding means for decoding a spectrum information code to spectrum information;

a noise level evaluating means for evaluating a noise level of a speech in a concerning decoding period by using a decoding result or the spectrum information code of at least one of spectrum information and power information, obtained from decoded spectrum information provided by the spectrum information decoding means, and outputting an evaluation result;

a first excitation codebook storing a plurality of non-noise time series vectors;

a second excitation codebook storing a plurality of noise time series

vectors;

an excitation codebook switching means for switching the first excitation codebook and the second excitation codebook based on the evaluation result of the noise level evaluating means;

a weighting-adding means for weighting the time series vectors from the first excitation codebook and second excitation codebook depending on their gains; and

a synthesis filter for producing a decoded speech based on an excitation signal, which is a weighted time series vector, and the decoded spectrum information from the spectrum information decoding means.

[Detailed Explanation of the Invention]

[0001]

[Technical Field of the Invention]

This invention relates to methods for speech coding and decoding and apparatuses for speech coding and decoding for performing compression coding of a speech signal to a digital signal. Particularly, this invention relates to a method for speech coding, method for speech decoding, apparatus for speech coding and apparatus for speech decoding for reproducing a high quality speech at low bit rates.

[0002]

[Related Art]

In the related art, code-excited linear prediction coding (Code-Excited Linear Prediction: CELP) coding is well-known as an efficient speech coding method, and its technique is described in "Code-excited linear prediction (CELP): High-quality speech at 8kbps" (ICASSP '85, pp. 937 – 940, by M. R.

Shroeder and B. S. Atal in 1985) .

[0003]

Fig. 6 illustrates an example of a whole configuration of a CELP speech coding and decoding method. In Fig. 6, an encoder 1, decoder 2, multiplexing means 3, and dividing means 4 are illustrated.

The encoder 1 includes a linear prediction parameter analyzing means 5, linear prediction parameter coding means 6, synthesis filter 7, adaptive codebook 8, excitation codebook 9, gain coding means 10, distance calculating means 11, and weighting-adding means 38. The decoder 2 includes a linear prediction parameter decoding means 12, synthesis filter 13, adaptive codebook 14, excitation codebook 15, gain decoding means 16, and weighting-adding means 39.

[0004]

In CELP speech coding, a speech in a frame of about 5 – 50 ms is divided into spectrum information and excitation information, and coded.

Explanations are made on operations in the CELP speech coding method. In the encoder 1, the linear prediction parameter analyzing means 5 analyzes an input speech S1, and extracts a linear prediction parameter, which is spectrum information of the speech. The linear prediction parameter coding means 6 codes the linear prediction parameter, and sets a coded linear prediction parameter as a coefficient for the synthesis filter 7.

[0005]

Explanations are made on coding of excitation information.

An old excitation signal is stored in the adaptive codebook 8. The adaptive codebook 8 outputs a time series vector, corresponding to an adaptive

code, which is generated by repeating the old excitation signal periodically.

A plurality of time series vectors trained by reducing a distortion between a speech for training and its coded speech for example is stored in the excitation codebook 9. The excitation codebook 9 outputs a time series vector corresponding to an excitation code.

Each of the time series vectors outputted from the adaptive codebook 8 and excitation codebook 9 is weighted by using a respective gain provided by the gain coding means 10 and added by the weighting-adding means 38. Then, an addition result is provided to the synthesis filter 7 as excitation signals, and a coded speech is produced. The distance calculating means 11 calculates a distance between the coded speech and the input speech S1, and searches an adaptive code, excitation code, and gains for minimizing the distance. When the above-stated coding is over, a linear prediction parameter code and the adaptive code, excitation code, and gain codes for minimizing a distortion between the input speech and the coded speech are outputted as a coding result.

[0006]

In the decoder 2, the linear prediction parameter decoding means 12 decodes the linear prediction parameter code to the linear prediction parameter, and sets the linear prediction parameter as a coefficient for the synthesis filter 13. The adaptive codebook 14 outputs a time series vector corresponding to an adaptive code, which is generated by repeating an old excitation signal periodically. The excitation codebook 15 outputs a time series vector corresponding to an excitation code. The time series vectors are weighted by using respective gains, which are decoded from the gain codes by

the gain decoding means 16, and added by the weighting-adding means 39. An addition result is provided to the synthesis filter 13 as an excitation signal, and an output speech S3 is produced.

[0007]

Among the CELP speech coding and decoding method, an improved speech coding and decoding method for reproducing a high quality speech according to the related art is described in "Phonetically – based vector excitation coding of speech at 3.6 kbps" (ICASSP '89, pp. 49 – 52, by S. Wang and A. Gersho in 1989).

Fig. 7 shows an example of a whole configuration of the speech coding and decoding method according to the related art, and same signs are used for means corresponding to the means in Fig. 6.

In Fig. 7, the encoder 1 includes a speech state deciding means 17, excitation codebook switching means 18, first excitation codebook 19, and second excitation codebook 20. The decoder 2 includes an excitation codebook switching means 21, first excitation codebook 22, and second excitation codebook 23.

Explanations are made on operations in the coding and decoding method in this configuration. In the encoder 1, the speech state deciding means 17 analyzes the input speech S1, and decides a state of the speech is which one of two states, e.g., voiced or unvoiced. The excitation codebook switching means 18 switches the excitation codebooks to be used in coding based on a speech state deciding result. For example, if the speech is voiced, the first excitation codebook 19 is used, and if the speech is unvoiced, the second excitation codebook 20 is used. Then, the excitation codebook

switching means 18 codes which excitation codebook is used in coding.

[0008]

In the decoder 2, the excitation codebook switching means 21 switches the first excitation codebook 22 and the second excitation codebook 23 based on a code showing which excitation codebook was used in the encoder 1, so that the excitation codebook, which was used in the encoder 1, is used in the decoder 2. According to this configuration, excitation codebooks suitable for coding in various speech states are provided, and the excitation codebooks are switched based on a state of an input speech. Hence, a high quality speech can be reproduced.

[0009]

A speech coding and decoding method of switching a plurality of excitation codebooks without increasing a transmission bit number according to the related art is disclosed in Japanese Unexamined Published Patent Application 8 – 185198. The plurality of excitation codebooks is switched based on a pitch frequency selected in an adaptive codebook, and an excitation codebook suitable for characteristics of an input speech can be used without increasing transmission data.

[0010]

[Problems to be Solved by the Invention]

As stated, in the speech coding and decoding method illustrated in Fig. 6 according to the related art, a single excitation codebook is used to produce a synthetic speech. Non-noise time series vectors with many pulses should be stored in the excitation codebook to produce a high quality coded speech even at low bit rates. Therefore, when a noise speech, e.g., background noise,

fricative consonant, etc., is coded and synthesized, there is a problem that a coded speech produces an unnatural sound, e.g., "Jiri-Jiri" and "Chiri-Chiri." This problem can be solved, if the excitation codebook includes only noise time series vectors. However, in that case, a quality of the coded speech degrades as a whole.

[0011]

In the improved speech coding and decoding method illustrated in Fig. 7 according to the related art, the plurality of excitation codebooks is switched based on the state of the input speech for producing a coded speech. Therefore, it is possible to use an excitation codebook including noise time series vectors in a noise period of the input speech and an excitation codebook including non-noise time series vectors in a voiced period other than the unvoiced noise period, for example. Hence, even if a noise speech is coded and synthesized, an unnatural sound, e.g., "Jiri-Jiri," is not produced. However, since the excitation codebook used in coding is also used in decoding, it becomes necessary to code and transmit data which excitation codebook was used. It becomes an obstacle for lowing bit rates.

[0012]

According to the speech coding and decoding method of switching the plurality of excitation codebooks without increasing a transmission bit number according to the related art, the excitation codebooks are switched based on a pitch period selected in the adaptive codebook. However, the pitch period selected in the adaptive codebook differs from an actual pitch period of a speech, and it is impossible to decide if a state of an input speech is noise or non-noise only from a value of the pitch period. Therefore, the problem that

the coded speech in the noise period of the speech is unnatural cannot be solved.

[0013]

This invention was intended to solve the above-stated problems. Particularly, this invention aims at providing speech coding and decoding methods for reproducing a high quality speech even at low bit rates.

[0014]

[Means to Solve the Problems]

In order to solve the above-stated problems, in a speech coding method according to this invention, a noise level of a speech in a concerning coding period is evaluated by using a code or coding result of at least one of spectrum information, power information, and pitch information, and a various excitation codebook is selected based on an evaluation result.

[0015]

In a speech coding method according to another invention, a plurality of excitation codebooks storing time series vectors with various noise levels is provided, and the plurality of excitation codebooks is switched based on an evaluation result of a noise level of a speech.

[0016]

In a speech coding method according to another invention, a noise level of time series vectors stored in an excitation codebook is changed based on an evaluation result of a noise level of a speech.

[0017]

In a speech coding method according to another invention, an excitation codebook storing noise time series vectors is provided. A low noise time series

vector is generated by sampling signal samples in the time series vectors based on the evaluation result of a noise level of a speech.

[0018]

In a speech coding method according to another invention, a first excitation codebook storing a noise time series vector and a second excitation codebook storing a non-noise time series vector are provided. A code vector is generated by adding the times series vector in the first excitation codebook and the time series vector in the second excitation codebook by weighting based on an evaluation result of a noise level of a speech.

[0019]

In a speech decoding method according to another invention, a noise level of a speech in a concerning decoding period is evaluated by using a code or coding result of at least one of spectrum information, power information, and pitch information, and a various excitation codebook is selected based on an evaluation result.

[0020]

In a speech decoding method according to another invention, a plurality of excitation codebooks storing time series vectors with various noise levels is provided, and the plurality of excitation codebooks is switched based on an evaluation result of the noise level of the speech.

[0021]

In a speech decoding method according to another invention, noise levels of time series vectors stored in excitation codebooks are changed based on an evaluation result of the noise level of the speech.

[0022]

In a speech decoding method according to another invention, an excitation codebook storing noise time series vectors is provided. A low noise time series vector is generated by sampling signal samples in the time series vectors based on the evaluation result of the noise level of the speech.

[0023]

In a speech decoding method according to another invention, a first excitation codebook storing a noise time series vector and a second excitation codebook storing a non-noise time series vector are provided. A code vector is generated by adding the times series vector in the first excitation codebook and the time series vector in the second excitation codebook by weighting based on an evaluation result of a noise level of a speech.

[0024]

A speech coding apparatus according to another invention includes a spectrum information encoding means for coding spectrum information of an input speech and outputting a coded spectrum information as an element of a coding result, a noise level evaluating means for evaluating a noise level of a speech in a concerning coding period by using a code or coding result of at least one of the spectrum information and power information, which is obtained from the coded spectrum information provided by the spectrum information encoding means, and outputting an evaluation result, a first excitation codebook storing a plurality of non-noise time series vectors, a second excitation codebook storing a plurality of noise time series vectors, an excitation codebook switching means for switching the first excitation codebook and the second excitation codebook based on the evaluation result by the noise level evaluating means, a weighting-adding means for weighting the time

series vectors from the first excitation codebook and second excitation codebook depending on their gains, a synthesis filter for producing a coded speech based on an excitation signal, which are weighted time series vectors, and the coded spectrum information provided by the spectrum information encoding means, and a distance calculating means for calculating a distance between the coded speech and the input speech, searching an excitation code and gain for minimizing the distance, and outputting a result as an excitation code, and a gain code as a coding result.

[0025]

A speech decoding apparatus according to another invention includes a spectrum information decoding means for decoding a spectrum information code to spectrum information, a noise level evaluating means for evaluating a noise level of a speech in a concerning decoding period by using a decoding result of at least one of the spectrum information and power information, which is obtained from decoded spectrum information provided by the spectrum information decoding means, and the spectrum information code and outputting an evaluating result, a first excitation codebook storing a plurality of non-noise time series vectors, a second excitation codebook storing a plurality of noise time series vectors, an excitation codebook switching means for switching the first excitation codebook and the second excitation codebook based on the evaluation result by the noise level evaluating means, a weighting-adding means for weighting the time series vectors from the first excitation codebook and the second excitation codebook depending on their gains, and a synthesis filter for producing a decoded speech based on an excitation signal, which is a weighted time series vector, and the decoded

spectrum information from the spectrum information decoding means.

[0026]

[Embodiments]

Explanations are made on embodiments of this invention with reference to drawings.

[0027]

Embodiment 1.

Same signs are used in Fig. 1 for corresponding elements of Fig. 7. Fig. 1 illustrates a whole configuration of a speech coding method and speech decoding method in embodiment 1 according to this invention. In Fig. 1, the linear prediction parameter analyzing means 5 is a spectrum information analyzing means for analyzing an input speech S1 and extracting a linear prediction parameter、which is spectrum information of the speech. The linear prediction parameter encoding means 6 is a spectrum information encoding means for coding the linear prediction parameter, which is the spectrum information and setting a coded linear prediction parameter as a coefficient for the synthesis filter 7. The first excitation codebooks 19 and 22 store pluralities of non-noise time series vectors, and the second excitation codebooks 20 and 23 store pluralities of noise time series vectors. The noise level evaluating means 24 and 26 evaluate a noise level, and the excitation codebook switching means 25 and 27 switch the excitation codebooks based on the noise level.

[0028]

Operations are explained.

In the encoder 1, the linear prediction parameter analyzing means 5

analyzes the input speech S1, and extracts a linear prediction parameter, which is spectrum information of the speech. The linear prediction parameter encoding means 6 codes the linear prediction parameter. Then, the linear prediction parameter encoding means 6 sets a coded linear prediction parameter as a coefficient for the synthesis filter 7, and also outputs the coded linear prediction parameter to the noise level evaluating means 24.

Explanations are made on coding of excitation information.

An old excitation signal is stored in the adaptive codebook 8, and a time series vector corresponding to an adaptive code, which is generated by repeating an old excitation signal periodically, is outputted. The noise level evaluating means 24 evaluates a noise level in a concerning coding period based on the coded linear prediction parameter inputted by the linear prediction parameter encoding means 6 and the adaptive code, e.g., a spectrum gradient, short-term prediction gain, and pitch fluctuation as shown in Fig. 2, and outputs an evaluation result to the excitation codebook switching means 25. The excitation codebook switching means 25 switches excitation codebooks for coding based on the evaluation result of the noise level. For example, if the noise level is low, the first excitation codebook 19 is used, and if the noise level is high, the second excitation codebook 20 is used.

[0029]

The first excitation codebook 19 stores a plurality of non-noise time series vectors, e.g., a plurality of time series vectors trained by reducing a distortion between a speech for training and its coded speech. The second excitation codebook 20 stores a plurality of noise time series vectors, e.g., a plurality of time series vectors generated from random noises. Each of the

first excitation codebook 19 and the second excitation codebook 20 outputs a time series vector respectively corresponding to an excitation code. Each of the time series vectors from the adaptive codebook 8 and one of first excitation codebook 19 or second excitation codebook 20 are weighted by using a respective gain provided by the gain encoder 10, and added by the weighting-adding means 38. An addition result is provided to the synthesis filter 7 as excitation signals, and a coded speech is produced. The distance calculating means 11 calculates a distance between the coded speech and the input speech S1, and searches an adaptive code, excitation code, and gain for minimizing the distance. When this coding is over, the linear prediction parameter code and an adaptive code, excitation code, and gain code for minimizing the distortion between the input speech and the coded speech are outputted as a coding result S2. These are characteristic operations in the speech coding method in embodiment 1.

[0030]

Explanations are made on the decoder 2. In the decoder 2, the linear prediction parameter decoding means 12 decodes the linear prediction parameter code to the linear prediction parameter, and sets the decoded linear prediction parameter as a coefficient for the synthesis filter 13, and outputs the decoded linear prediction parameter to the noise level evaluating means 26.

Explanations are made on decoding of excitation information. The adaptive codebook 14 outputs a time series vector corresponding to an adaptive code, which is generated by repeating an old excitation signal periodically. The noise level evaluating means 26 evaluates a noise level by using the decoded linear prediction parameter inputted by the linear prediction

parameter decoding means 12 and the adaptive code in a same method with the noise level evaluating means 24 in the encoder 1, and outputs an evaluation result to the excitation codebook switching means 27. The excitation codebook switching means 27 switches the first excitation codebook 22 and the second excitation codebook 23 based on the evaluation result of the noise level in a same method with the excitation codebook switching means 25 in the encoder 1.

[0031]

A plurality of non-noise time series vectors, e.g., a plurality of time series vectors generated by training for reducing a distortion between a speech for training and its coded speech, is stored in the first excitation codebook 22. A plurality of noise time series vectors, e.g., a plurality of vectors generated from random noises, is stored in the second excitation codebook 23. Each of the first and second excitation codebooks outputs a time series vector respectively corresponding to an excitation code. The time series vectors from the adaptive codebook 14 and one of excitation codebook 22 or excitation codebook 23 are weighted by using respective gains, decoded from gain codes by the gain decoding means 16, and added by the weighting-adding means 39. An addition result is provided to the synthesis filter 13 as an excitation signal, and an output speech S3 is produced. These are operations are characteristic operations in the speech decoding method in embodiment 1.

[0032]

In embodiment 1, the noise level of the input speech is evaluated by using the code and coding result, and various excitation codebooks are used based on the evaluation result. Therefore, a high quality speech can be

reproduced with a small data amount.

[0033]

Embodiment 2.

In embodiment 1, two excitation codebooks are switched. However, it is also possible that three or more excitation codebooks are provided and switched based on a noise level.

In embodiment 2, a suitable excitation codebook can be used even for a medium speech, e.g., slightly noisy, in addition to two kinds of speech, i.e., noise and non-noise. Therefore, a high quality speech can be reproduced.

[0034]

Embodiment 3.

Fig. 3 shows a whole configuration of a speech coding method and speech decoding method in embodiment 3 of this invention. In Fig. 3, same signs are used for units corresponding to the units in Fig. 1. In Fig. 3, excitation codebooks 28 and 30 store noise time series vectors, and sampling means 29 and 31 set an amplitude value of a sample with a low amplitude in the time series vectors to zero.

[0035]

Operations are explained. In the encoder 1, the linear prediction parameter analyzing means 5 analyzes the input speech S1, and extracts a linear prediction parameter, which is spectrum information of the speech. The linear prediction parameter encoding means 6 codes the linear prediction parameter. Then, the linear prediction parameter encoding means 6 sets a coded linear prediction parameter as a coefficient for the synthesis filter 7, and also outputs the coded linear prediction parameter to the noise level evaluating

means 24.

Explanations are made on coding of excitation information. An old excitation signal is stored in the adaptive codebook 8, and a time series vector corresponding to an adaptive code, which is generated by repeating an old excitation signal periodically, is outputted. The noise level evaluating means 24 evaluates a noise level in a concerning coding period by using the coded linear prediction parameter, which is inputted from the linear prediction parameter encoding means 6, and an adaptive code, e.g., a spectrum gradient, short-term prediction gain, and pitch fluctuation, and outputs an evaluation result to the sampling means 29.

[0036]

The excitation codebook 28 stores a plurality of time series vectors generated from random noises, for example, and outputs a time series vector corresponding to an excitation code. If the noise level is low in the evaluation result of the noise, the sampling means 29 outputs a time series vector, in which an amplitude of a sample with an amplitude below a determined value in the time series vectors, inputted from the excitation codebook 28, is set to zero, for example. If the noise level is high, the sampling means 29 outputs the time series vector inputted from the excitation codebook 28 without modification. Each of the times series vectors from the adaptive codebook 8 and the sampling means 29 is weighted by using a respective gain provided by the gain encoder 10 and added by the weighting-adding means 38. An addition result is provided to the synthesis filter 7 as excitation signals, and a coded speech is produced. The distance calculating means 11 calculates a distance between the coded speech and the input speech S1, and searches an

adaptive code, excitation code, and gain for minimizing the distance. When coding is over, the linear prediction parameter code and the adaptive code, excitation code, and gain code for minimizing a distortion between the input speech and the coded speech are outputted as a coding result S2. These are characteristic operations in the speech coding method in embodiment 3.

[0037]

Explanations are made on the decoder 2. In the decoder 2, the linear prediction parameter decoding means 12 decodes the linear prediction parameter code to the linear prediction parameter. The linear prediction parameter decoding means 12 sets the linear prediction parameter as a coefficient for the synthesis filter 13, and also outputs the linear prediction parameter to the noise level evaluating means 26.

Explanations are made on decoding of excitation information. The adaptive codebook 14 outputs a time series vector corresponding to an adaptive code, generated by repeating an old excitation signal periodically. The noise level evaluating means 26 evaluates a noise level by using the decoded linear prediction parameter inputted from the linear prediction parameter decoding means 12 and the adaptive code in a same method with the noise level evaluating means 24 in the encoder 1, and outputs an evaluation result to the sampling means 31.

[0038]

The excitation codebook 30 outputs a time series vector corresponding to an excitation code. The sampling means 31 outputs a time series vector based on the evaluation result of the noise level in same processing with the sampling means 29 in the encoder 1. Each of the time series vectors

outputted from the adaptive codebook 14 and sampling means 31 are weighted by using a respective gain provided by the gain decoder 16, and added by the weighting-adding means 39. An addition result is provided to the synthesis filter 13 as an excitation signal, and an output speech S3 is produced.

[0039]

In embodiment 3, the excitation codebook storing noise time series vectors is provided, and an excitation with a low noise level can be generated by sampling excitation signal samples based on an evaluation result of the noise level the speech. Hence, a high quality speech can be reproduced with a small data amount. Further, since it is not necessary to provide a plurality of excitation codebooks, a memory amount for storing the excitation codebook can be reduced.

[0040]

Embodiment 4.

In embodiment 3, the samples in the time series vectors are either sampled or not. However, it is also possible to change a threshold value of an amplitude for sampling the samples based on the noise level. In embodiment 4, a suitable time series vector can be generated and used also for a medium speech, e.g., slightly noisy, in addition to the two types of speech, i.e., noise and non-noise. Therefore, a high quality speech can be reproduced.

[0041]

Embodiment 5.

Fig. 4 shows a whole configuration of a speech coding method and a speech decoding method in embodiment 5 of this invention, and same signs are used for units corresponding to the units in Fig. 1.

In Fig. 4, first excitation codebooks 32 and 35 store noise time series vectors, and second excitation codebooks 33 and 36 store non-noise time series vectors. The weight determining means 34 and 37 are also illustrated.

[0042]

Operations are explained. In the encoder 1, the linear prediction parameter analyzing means 5 analyzes the input speech S1, and extracts a linear prediction parameter, which is spectrum information of the speech. The linear prediction parameter encoding means 6 codes the linear prediction parameter. Then, the linear prediction parameter encoding means 6 sets a coded linear prediction parameter as a coefficient for the synthesis filter 7, and also outputs the coded prediction parameter to the noise level evaluating means 24.

Explanations are made on coding of excitation information. The adaptive codebook 8 stores an old excitation signal, and outputs a time series vector corresponding to an adaptive code, which is generated by repeating an old excitation signal periodically. The noise level evaluating means 24 evaluates a noise level in a concerning coding period by using the coded linear prediction parameter, which is inputted from the linear prediction parameter encoding means 6 and the adaptive code, e.g., a spectrum gradient, short-term prediction gain, and pitch fluctuation, and outputs an evaluation result to the weight determining means 34.

[0043]

The first excitation codebook 32 stores a plurality of noise time series vectors generated from random noises, for example, and outputs a time series vector corresponding to an excitation code. The second excitation codebook 33

stores a plurality of time series vectors generated by training for reducing a distortion between a speech for training and its coded speech, and outputs a time series vector corresponding to an excitation code. The weight determining means 34 determines a weight provided to the time series vector from the first excitation codebook 32 and the time series vector from the second excitation codebook 33 based on the evaluation result of the noise level inputted from the noise level evaluating means 24, as illustrated in Fig. 5, for example. Each of the time series vectors from the first excitation codebook 32 and the second excitation codebook 33 is weighted by using the weight provided by the weight determining means 34, and added. The time series vector outputted from the adaptive codebook 8 and the time series vector, which is generated by being weighted and added, are weighted by using respective gains provided by the gain encoding means 10, and added by the weighting-adding means 38. Then, an addition result is provided to the synthesis filter 7 as excitation signals, and a coded speech is produced. The distance calculating means 11 calculates a distance between the coded speech and the input speech S1, and searches an adaptive code, excitation code, and gain for minimizing the distance. When coding is over, the linear prediction parameter code, adaptive code, excitation code, and gain code for minimizing a distortion between the input speech and the coded speech, are outputted as a coding result.

[0044]

Explanations are made on the decoder 2. In the decoder 2, the linear prediction parameter decoding means 12 decodes the linear prediction parameter code to the linear prediction parameter. Then, the linear

prediction parameter decoding means 12 sets the linear prediction parameter as a coefficient for the synthesis filter 13, and also outputs the linear prediction parameter to the noise evaluating means 26.

Explanations are made on decoding of excitation information. The adaptive codebook 14 outputs a time series vector corresponding to an adaptive code by repeating an old excitation signal periodically. The noise level evaluating means 26 evaluates a noise level by using the decoded linear prediction parameter, which is inputted from the linear prediction parameter decoding means 12, and the adaptive code in a same method with the noise level evaluating means 24 in the encoder 1, and outputs an evaluation result to the weight determining means 37.

[0045]

The first excitation codebook 35 and the second excitation codebook 36 output time series vectors corresponding to excitation codes. The weight determining means 37 weights based on the noise level evaluation result inputted from the noise level evaluating means 26 in a same method with the weight determining means 34 in the encoder 1. Each of the time series vectors from the first excitation codebook 35 and the second excitation codebook 36 is weighted by using a respective weight provided by the weight determining means 37, and added. The time series vector outputted from the adaptive codebook 14 and the time series vector, which is generated by being weighted and added, are weighted by using respective gains decoded from the gain codes by the gain decoding means 16, and added by the weighting-adding means 39. Then, an addition result is provided to the synthesis filter 13 as an excitation signal, and an output speech S3 is produced.

[0046]

In embodiment 5, the noise level of the speech is evaluated by using a code and coding result, and the noise time series vector or non-noise time series vector are weighted based on the evaluation result, and added. Therefore, a high quality speech can be reproduced with a small data amount.

[0047]

Embodiment 6.

In embodiments 1 – 5, it is also possible to change gain codebooks based on the evaluation result of the noise level. In embodiment 6, a most suitable gain codebook can be used based on the excitation codebook. Therefore, a high quality speech can be reproduced.

[0048]

Embodiment 7.

In embodiments 1 – 6, the noise level of the speech is evaluated, and the excitation codebooks are switched based on the evaluation result. However, it is also possible to decide and evaluate each of a voiced onset, plosive consonant, etc., and switch the excitation codebooks based on an evaluation result. In embodiment 7, in addition to the noise state of the speech, the speech is classified in more details, e.g., voiced onset, plosive consonant, etc., and a suitable excitation codebook can be used for each state. Therefore, a high quality speech can be reproduced.

[0049]

[Effects of the Invention]

In the speech coding method, speech decoding method, speech coding apparatus, and speech decoding apparatus according to this invention, a noise

level of a speech in a concerning coding period is evaluated by using a code or coding result of at least one of the spectrum information, power information, and pitch information, and various excitation codebooks are used based on the evaluation result. Therefore, a high quality speech can be reproduced with a small data amount.

[0050]

In the speech coding method and speech decoding method according to this invention, a plurality of excitation codebooks storing excitations with various noise levels is provided, and the plurality of excitation codebooks is switched based on the evaluation result of the noise level of the speech. Therefore, a high quality speech can be reproduced with a small data amount.

[0051]

In the speech coding method and speech decoding method according to this invention, the noise levels of the time series vectors stored in the excitation codebooks are changed based on the evaluation result of the noise level of the speech. Therefore, a high quality speech can be reproduced with a small data amount.

[0052]

In the speech coding method and speech decoding method according to this invention, an excitation codebook storing noise time series vectors is provided, and a time series vector with a low noise level is generated by sampling signal samples in the time series vectors based on the evaluation result of the noise level of the speech. Therefore, a high quality speech can be reproduced with a small data amount.

[0053]

In the speech coding method and speech decoding method according to this invention, the first excitation codebook storing noise time series vectors and the second excitation codebook storing non-noise time series vectors are provided, and the time series vector in the first excitation codebook or the time series vector in the second excitation codebook is weighted based on the evaluation result of the noise level of the speech, and added to generate a time series vector. Therefore, a high quality speech can be reproduced with a small data amount.

[Brief Description of the Drawings]

[Fig. 1] Fig. 1 shows a block diagram of a whole configuration of a speech coding and speech decoding apparatus in embodiment 1 of this invention.

[Fig. 2] Fig. 2 shows a table for explaining an evaluation of a noise level in embodiment 1 of this invention illustrated in Fig. 1.

[Fig. 3] Fig. 3 shows a block diagram of a whole configuration of a speech coding and speech decoding apparatus in embodiment 3 of this invention.

[Fig. 4] Fig. 4 shows a block diagram of a whole configuration of a speech coding and speech decoding apparatus in embodiment 5 of this invention.

[Fig. 5] Fig. 5 shows a schematic line chart for explaining a decision process of weighting in embodiment 5 illustrated in Fig. 4.

[Fig. 6] Fig. 6 shows a block diagram of a whole configuration of a CELP speech coding and decoding apparatus according to the related art.

[Fig. 7] Fig. 7 shows a block diagram of a whole configuration of an

improved CELP speech coding and decoding apparatus according to the related art.

[Explanations of Signs]

1: encoder, 2: decoder, 3: multiplexing means, 4: dividing means, 5: linear prediction parameter analyzing means, 6: linear prediction parameter coding means, 7, 13: synthesis filter, 8, 14: adaptive codebook, 9, 15: excitation codebook, 10: gain coding means, 11: distance calculating means, 12: linear prediction parameter decoding means, 16: gain decoding means, 17: speech state deciding means, 18, 21: excitation codebook switching means, 19, 22: first excitation codebook; 20, 23: second excitation codebook, 24, 26: excitation codebook switching means, 25, 27: excitation codebook switching means 28, 30: excitation codebook, 29, 31: sampling means, 32, 35: first excitation codebook 33, 36: second excitation codebook, 34, 37: weight determining means, 38, 39: weighting-adding means

[Document Name] abstract

[Abstract]

[Aim] A high quality speech is reproduced with a small data amount in speech coding and decoding for performing compression coding and decoding of a speech signal to a digital signal.

[Means to Solve] In speech coding method according to a code-excited linear prediction (CELP) speech coding, a noise level of a speech in a concerning coding period is evaluated by using a code or coding result of at least one of spectrum information, power information, and pitch information, and various excitation codebooks are used based on an evaluation result

[Selected Figure] Fig. 1

Fig.1

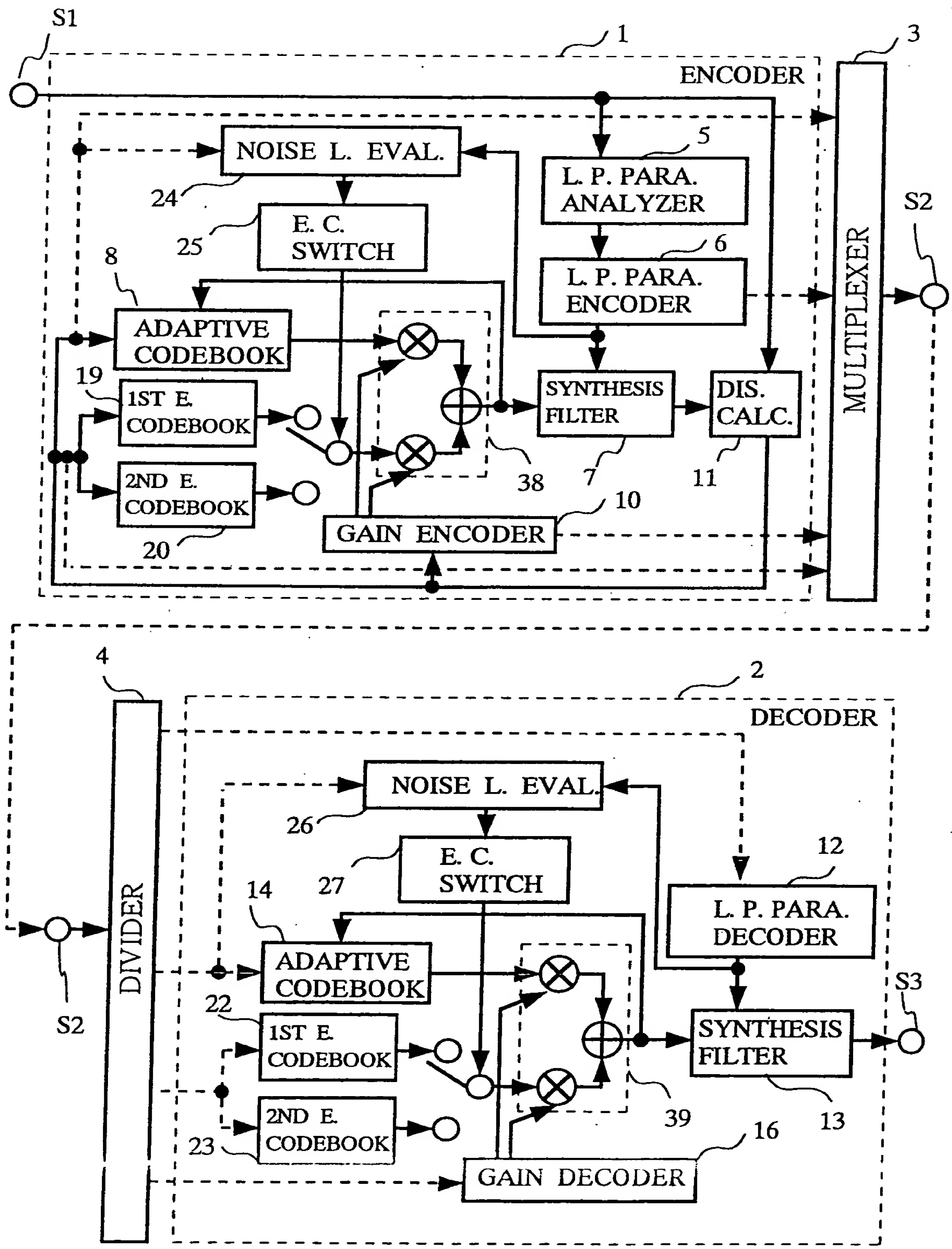
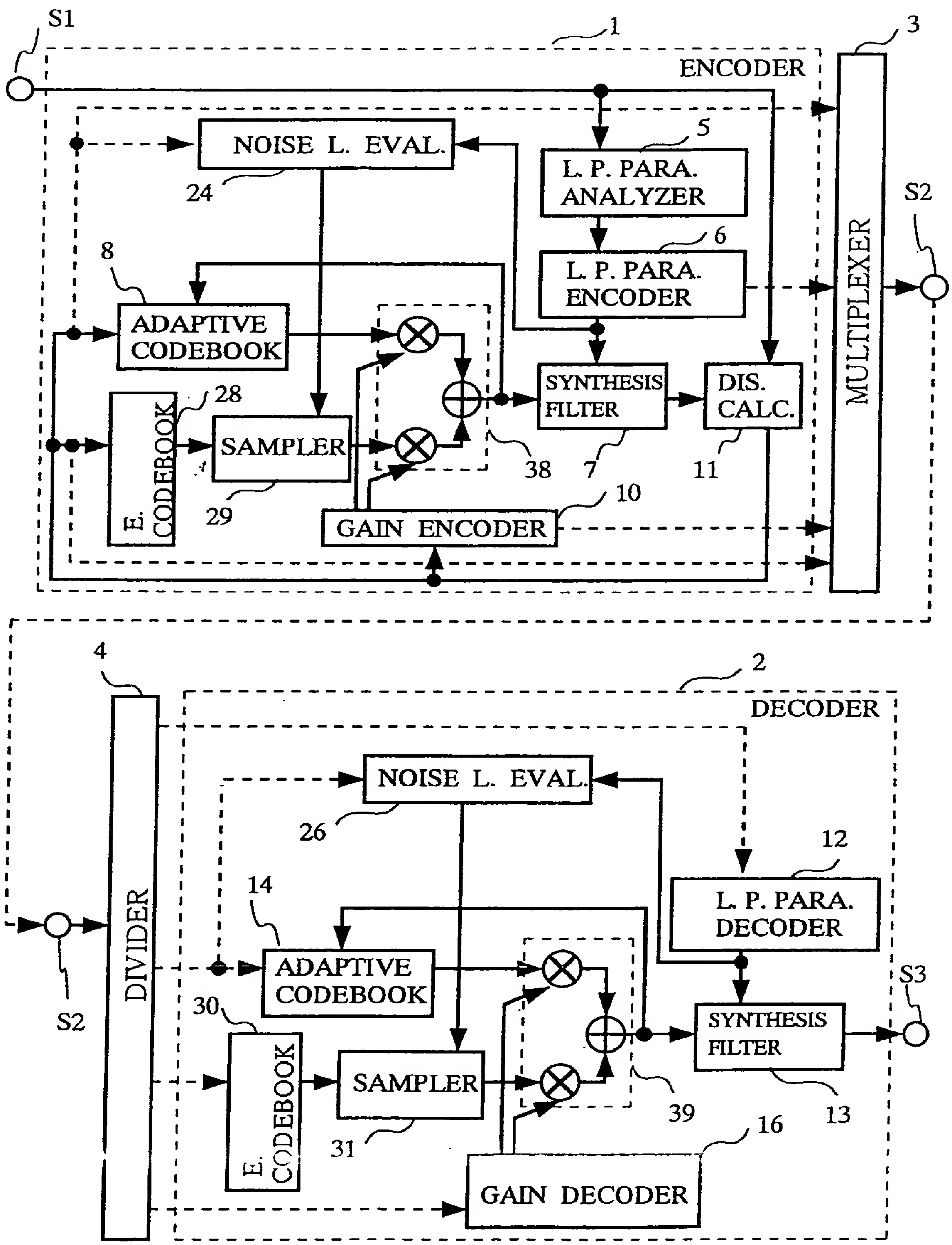

S1
1
ENCODER
3
NOISE L. EVAL.
5
L. P. PARA.
ANALYZER
24
25
E. C.
SWITCH
S2
8
6
L. P. PARA.
ENCODER
MULTIPLEXER
ADAPTIVE
CODEBOOK
19
1ST E.
CODEBOOK
SYNTHESIS
FILTER
DIS.
CALC.
2ND E.
CODEBOOK
38
7
11
10
GAIN ENCODER
20
4
2
DECODER
NOISE L. EVAL.
26
E. C.
SWITCH
12
14
27
L. P. PARA.
DECODER
DIVIDER
ADAPTIVE
CODEBOOK
S3
22
1ST E.
CODEBOOK
SYNTHESIS
FILTER
S2
2ND E.
CODEBOOK
39
13
16
23
GAIN DECODER

Fig.2

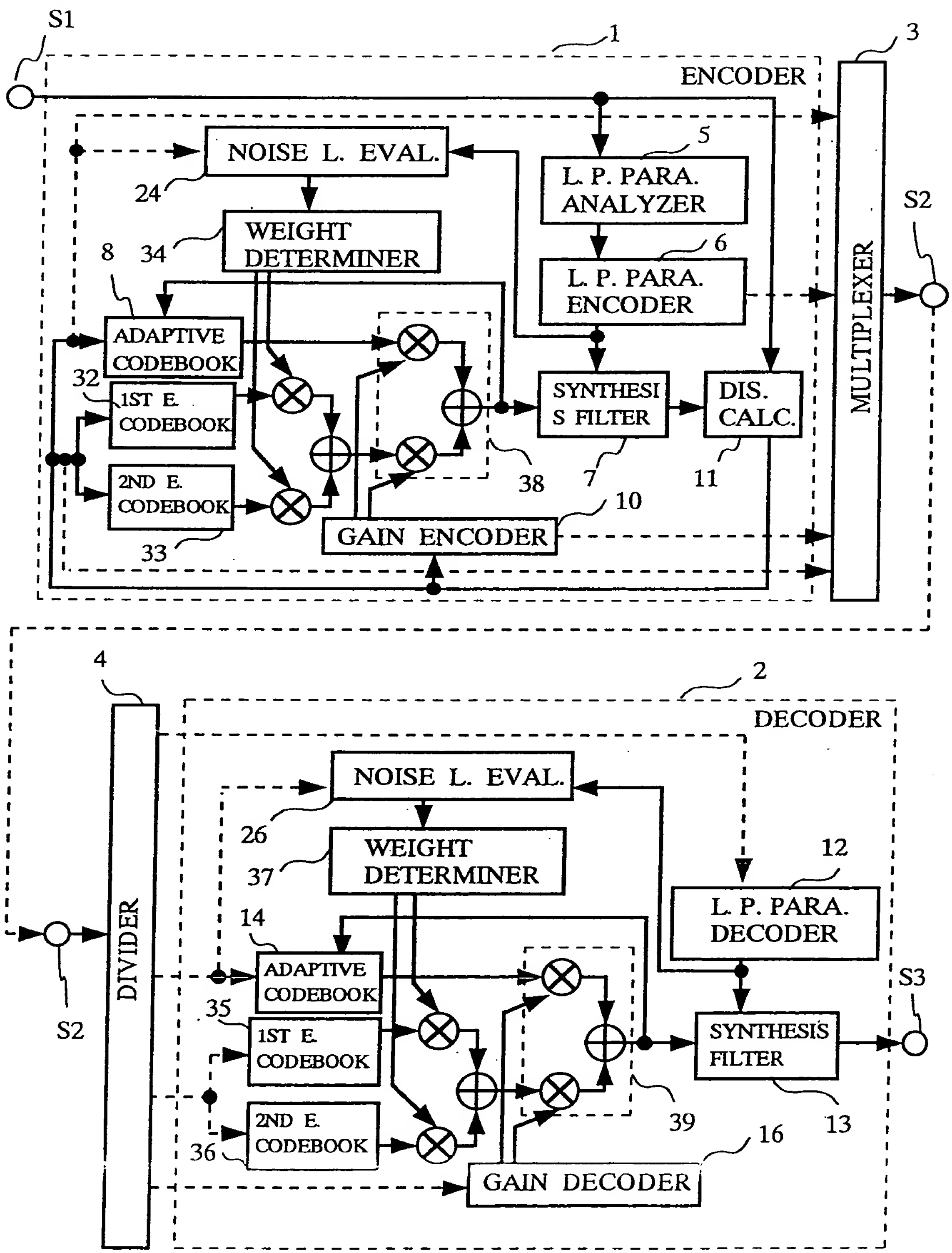
NOISE LEVEL	S ⟷ L
SPECTRUM GRADIENT	LOW GRADIENT ⟷ FLAT, HIGH GRADIENT
SHORT-TERM PREDICTION GAIN	L ⟷ S
PITCH FLUCTUATION	S ⟷ L

Fig.3



S1
1
ENCODER
3
NOISE L. EVAL.
5
L. P. PARA. ANALYZER
24
S2
8
6
L. P. PARA. ENCODER
ADAPTIVE CODEBOOK
MULTIPLEXER
SYNTHESIS FILTER
DIS. CALC.
28
E. CODEBOOK
SAMPLER
38
7
11
29
10
GAIN ENCODER
4
2
DECODER
NOISE L. EVAL.
26
12
14
L. P. PARA. DECODER
DIVIDER
ADAPTIVE CODEBOOK
S3
30
S2
SYNTHESIS FILTER
E. CODEBOOK
SAMPLER
39
13
31
16
GAIN DECODER

Fig.4



S1
1
3
ENCODER
NOISE L. EVAL.
5
L. P. PARA. ANALYZER
24
WEIGHT DETERMINER
8
34
6
L. P. PARA. ENCODER
MULTIPLEXER
S2
ADAPTIVE CODEBOOK
32
1ST E. CODEBOOK
SYNTHESIS FILTER
DIS. CALC.
2ND E. CODEBOOK
38
7
11
10
GAIN ENCODER
33
4
2
DECODER
NOISE L. EVAL.
26
WEIGHT DETERMINER
37
12
L. P. PARA. DECODER
14
ADAPTIVE CODEBOOK
S3
35
1ST E. CODEBOOK
SYNTHESIS FILTER
S2
DIVIDER
2ND E. CODEBOOK
39
13
36
16
GAIN DECODER

Fig.5

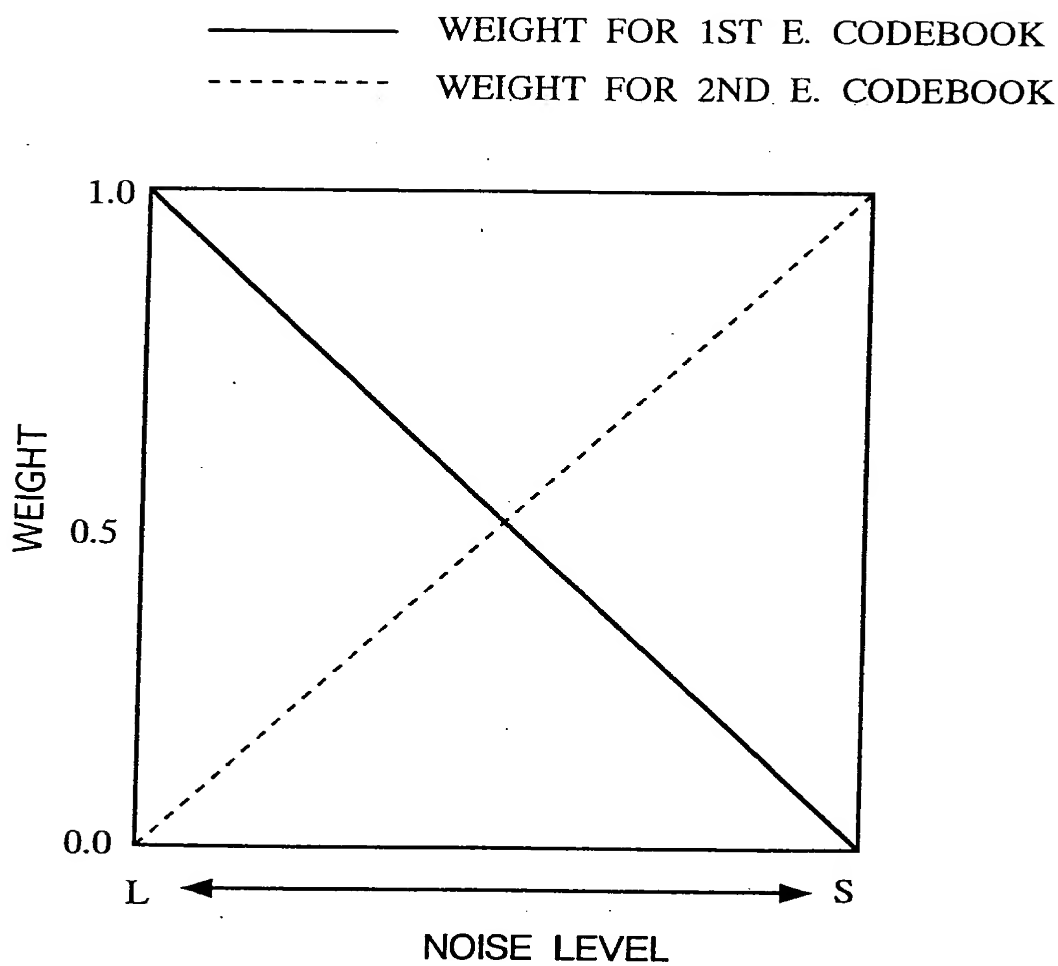

WEIGHT FOR 1ST E. CODEBOOK
WEIGHT FOR 2ND E. CODEBOOK
1.0
WEIGHT
0.5
0.0
L
S
NOISE LEVEL

Fig.6

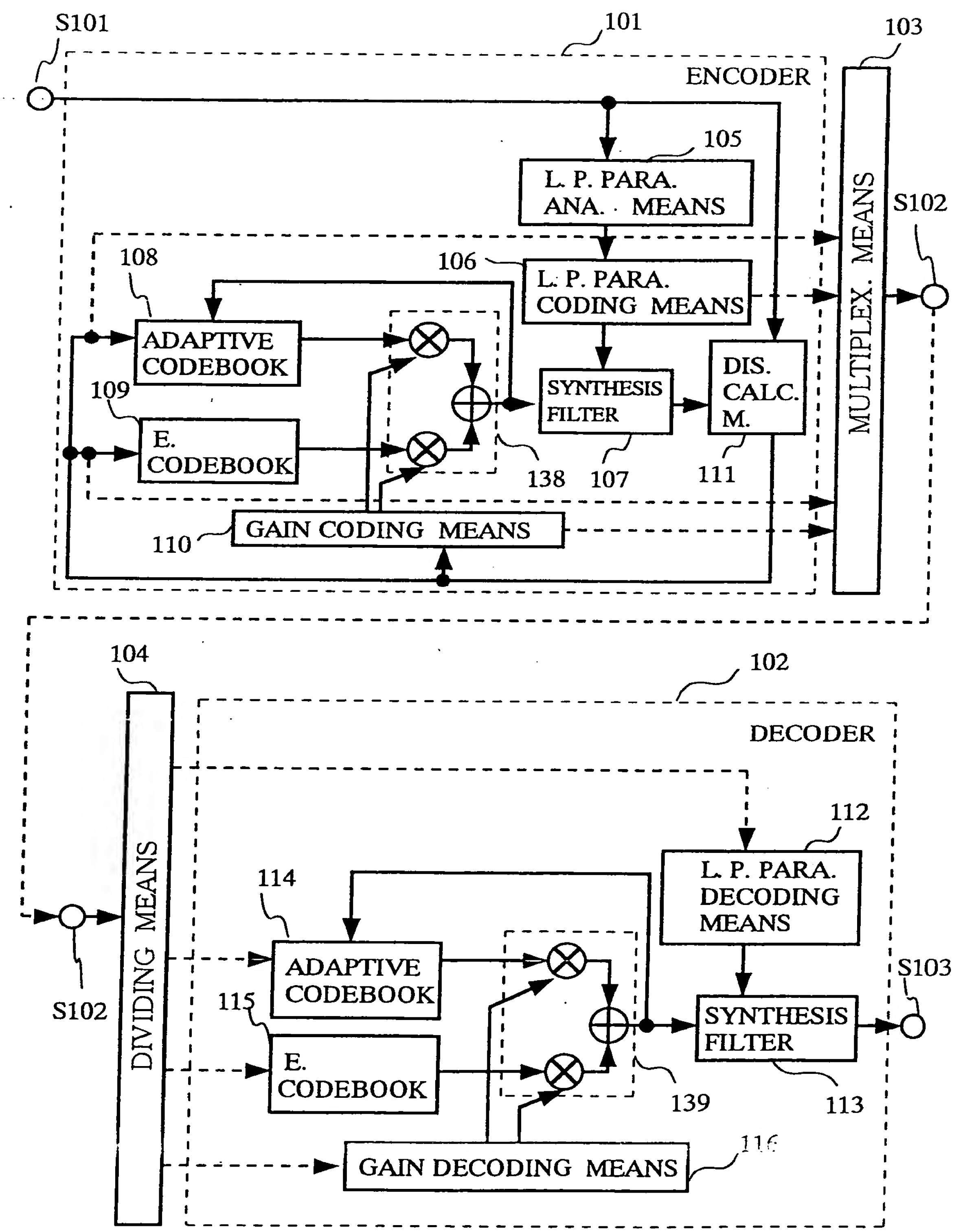
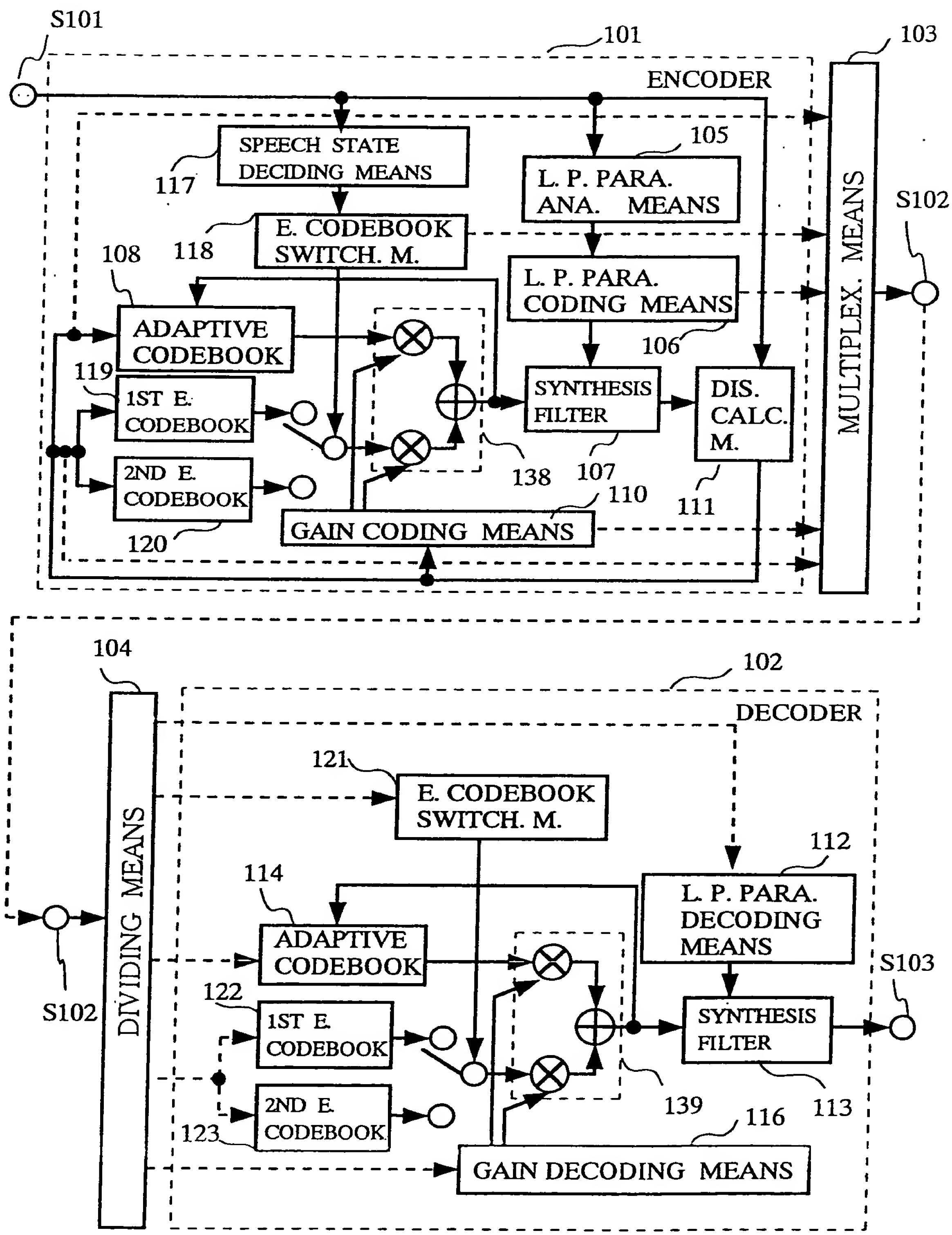

S101
101
103
ENCODER
105
L. P. PARA.
ANA. MEANS
S102
108
106
L. P. PARA.
CODING MEANS
ADAPTIVE
CODEBOOK
DIS.
CALC.
M.
109
SYNTHESIS
FILTER
MULTIPLEX. MEANS
E.
CODEBOOK
138
107
111
GAIN CODING MEANS
110
104
102
DECODER
112
L. P. PARA.
DECODING
MEANS
114
DIVIDING MEANS
ADAPTIVE
CODEBOOK
S103
115
SYNTHESIS
FILTER
S102
E.
CODEBOOK
139
113
GAIN DECODING MEANS
116

Fig.7


S101
101
103
ENCODER
SPEECH STATE DECIDING MEANS
105
L. P. PARA. ANA. MEANS
117
E. CODEBOOK SWITCH. M.
S102
108
118
L. P. PARA. CODING MEANS
MULTIPLEX. MEANS
ADAPTIVE CODEBOOK
106
119
1ST E. CODEBOOK
SYNTHESIS FILTER
DIS. CALC. M.
2ND E. CODEBOOK
138
107
110
111
GAIN CODING MEANS
120
104
102
DECODER
121
E. CODEBOOK SWITCH. M.
112
114
L. P. PARA. DECODING MEANS
ADAPTIVE CODEBOOK
DIVIDING MEANS
S103
122
1ST E. CODEBOOK
SYNTHESIS FILTER
S102
2ND E. CODEBOOK
139
116
113
123
GAIN DECODING MEANS